<u>**Final – Jan 2003 to Dec 2004**</u>

Steganalysis Techniques for Documents and Images
N. Memon, Ed Wong and Xiaolin Wu

1. **Cover Sheet**: Attached.

2. **Objectives:** The objective of this project was to develop novel techniques and algorithms for document and image steganalysis.

3. **Status of effort:** We have met and even exceeded our goals that we had outlined for our project. Our efforts have focused on two areas. The first area is LSB steganalysis for grayscale images. Here, as we had proposed (as a challenging task), we have generalized our previous steganalysis technique of sample pair analysis to a theoretical framework for the detection of the LSB steganography. The new framework exploits high-order statistics of the samples, and sheds some light on the effects of window size in steganalysis, an issue not well understood before.

The second thrust of our efforts has been on steganalysis of binary documents. Last year we developed steganalysis methods to detect marked images hidden by following data hiding techniques:

1. Boundary-based approaches
2. The Line, word, or character shifting
3. Fixed partitioning of image
4. Modification of character features
5. Modification of run-lengths
6. Modification of half-tone patterns

In addition to electronic documents we have also developed steganalysis methods to detect marked images when the original cover images are scanned/printed images. It is known that this is one of the most challenging problems in steganalysis literature. We have made significant progress and our work has laid the foundation to the goal of successfully detecting messages hidden in scanned/printed images. It should be noted that this goal was not part of the proposed work.

4. **Accomplishments/New Findings:** Below we summarize the main accomplishments of our project:

a. An exciting progress made is the generalization of our previous steganalysis technique of sample pair analysis to a theoretical framework for the detection of the LSB steganography. The new framework exploits high-order statistics of the samples, and sheds some light on the effects of window size in steganalysis, an issue not well understood before. We have developed new detection algorithms

20050608 055

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruc
the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other
reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis High
Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 1	3. REPORT TYPE AND DATES COVERED Final Report	
4. TITLE AND SUBTITLE Steganalysis Techniques for LSB Embedding in Documents and Images			**5. FUNDING NUMBERS** F49620-03-1-0066
6. AUTHOR(S) N. Memon, Ed Wong and Xiaolin Wu			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Polytechnic University Six Metrotech Center Brooklyn, NY 11201			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited	

13. ABSTRACT *(Maximum 200 Words)*

The objective of the proposed research is to develop novel techniques and algorithms for document and image steganalysis. We developed a new framework for LSB steganalysis that gives better performance then the best of the known techniques. We also developed a new methodology of steganalysis by modeling steganography as an additive noise process and measuring the mean and variance of the stego-signal. We derived a formula that governs all additive steganography. We then illustrated the efficacy of our model using variations of LSB steganography. For binary images, we have made significant progress in developing steganalysis methods to detect marked images when the original cover images are either purely electronic images or printed-scanned images. Our work represents the first effort for detecting steganography in binary images.

14. SUBJECT TERMS			15. NUMBER OF PAGES 4
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

in this new framework. Preliminary simulation results point to improved estimation of hidden message length in continuous-tone images compared to the best of existing techniques (RS Steganalysis). We proposed a general framework for the detection of the LSB steganography using digital media files as cover objects. The new framework exploits high-order statistics of the samples. It can compute a robust estimate of the length of a secret message hidden in the least significant bits of samples for a large class of digital media contents such as image, video and audio, in which the underlying signals consist of correlated samples. A case study on the LSB steganalysis of natural grey-scale and color images and experimental results was done.

b. We developed a new methodology for steganalysis by modeling steganography as an additive noise process and measuring the mean and variance of the stego-signal. We derived a formula that governs all additive steganography. It can be applied to any steganographic technique if it the embedding process can be modeled as an additive noise process. We then used steganalysis of LSB embedding as a case study, and we achieved performance that is comparable to the most advanced LSB steganalysis techniques in the literature. We also showed that our framework performs effectively for other steganographic techniques, for example, +- embedding, and boundary-based embedding in binary images.

c. Our second accomplishment was the steganalysis of pixel flipping embedding using compressed bit rate. Pixel flipping embedding refers to the binary document steganography techniques based on boundary modifications, fixed partitioning of the image into blocks, and modification of run-length patterns. All these techniques share a common characteristic - they embed information by flipping image pixels. We observed that embedding increases the compressed bit rate. Using JBIG2 compression algorithm, we empirically derived a function between flipping rate and compressed bit rate. This allowed us to estimate message length accurately.

d. The third accomplishment was the steganalysis of halftone images. We developed a steganalysis method of halftone images without knowledge of the original cover object. We first convert halftone images into grayscale like images by low-pass filtering. Then a set of features are extracted from the filtered images by decomposing with quadrature mirror filters at different scale and orientation. Finally statistical training is performed to catch the "unusual" properties of marked images. Experimental results demonstrate that the proposed approach is effective and accurate.

e. The fourth accomplishment was the steganalysis of degraded documents. We considered the case when naturally degraded images, such as printed and scanned images, were used as cover images. When data were hidden in printed/scanned images, the embedding noise would be carefully covered by the printing/scanning noise. To do steganalysis, we used a nonlinear document degradation model to show that printing/scanning and embedding introduce different types of noises.

Then a practical approach was developed to detect marked images. Experiments gave acceptable accuracy.

5. Personnel Supported:
X. Wu, E. K. Wong, N. Memon, and M. Jiang.

6. Publications:
1. S. Dumitrescu, X. Wu and N. Memon, "On steganalysis of random LSB embedding in continuous-tone images", *The Proc. of the 2002 International Conference on Image Processing*, pp. 641-644, 2002.
2. S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSB Steganography via Sample Pair Analysis", *IEEE Trans. on Signal Processing*, vol. 51, no. 7, July, pp. 1995-2007, 2003.
3. M. Jiang, X. Wu, E. K. Wong, and N. Memon, "Steganalysis of boundary-based steganography using autoregressive model of digital boundaries," in *IEEE Int'l Conference on Multimedia and Expo*, Taipei, Taiwan, June 2004.
4. M. Jiang, N. Memon, E. K. Wong, and X. Wu, "Quantitative steganalysis of binary images," in Proceedings of the *IEEE Int'l Conference on Image Processing*, Singapore, Oct. 2004.
5. M. Jiang, E. K. Wong, N. Memon ,and X. Wu, "A simple technique for estimating message lengths for additive noise steganography," in Proceedings of *The Eighth International Conference on Control, Automation, Robotics and Vision*, Kungming, China, Dec. 2004, accepted.
6. M. Jiang, E. K. Wong, N. Memon, and X. Wu, "Universal Steganalysis of Binary Images Using Pattern Variation," in preparation for *Journal of Electronic Imaging*.
7. M. Jiang, E. K. Wong, N. Memon ,and X. Wu, "Steganalysis of Halftone Images", in preparation for *IEEE Transactions on Information Forensics and Security*.
8. M. Jiang, E. K. Wong, N. Memon ,and X. Wu, "Estimating Message Lengths for Additive Noise Steganography," in preparation for *IEEE Transactions on Image Processing*.
9. M. Jiang, E. K. Wong, N. Memon ,and X. Wu, "An Improved Steganalysis Technique of Binary Images," in preparation for *IEEE Transactions on Image Processing*.
10. M. Jiang, E. K. Wong, N. Memon, and X. Wu, "Steganalysis of Halftone Images," *IEEE Int'l Conf on Acoustics, Speech, and Signal Processing*, March 18-23, 2005, Philadelphia, PA.
11. M. Jiang, E. K. Wong, N. Memon, and X. Wu, *"Steganalysis of Printed/Scanned Document Images,"* to be submitted to IEEE International Workshop on Multimedia Signal Processing, Shanghai, Oct 30 - Nov 2, 2005.
12. M. Chen, N. Memon, and E. K. Wong, "Data hiding in document images," Multimedia Security*: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property*, (edited by C. S. Lu), Idea Group Inc., Hershey, PA., 2004.
13. S. Dumitrescu, X. Wu, "LSB Steganalysis Based on High-order Statistics," to appear *in Proceedings of ACM Multimedia Security Workshop 2005*, August 2005.

14. S. Dumitrescu and X. Wu, "A new framework of LSB steganalysis of digital media", to appear in *IEEE Transactions on Signal Processing, Supplement on Secure Media*, Oct. 2005.
15. M. Jiang. Steganography and Steganalysis for Binary Documents. PhD thesis. To be completed August 2005.

7. Interactions/Transitions:
a. Attended and presented papers at SPIE Watermarking and Multimedia Security Conference, San Jose, 2003.
b. Presentation to visitors from AFRL at site meeting on July 29, 2003.
c. Attended and presented papers at SPIE Watermarking and Multimedia Security Conference, San Jose, 2004.
d. Presentation at Air Force Office of Scientific Research PI Conference, Cornell, July 2004.
e. Attended IEEE Int'l Conference on Multimedia and Expo, Taipei, Taiwan, June 2004.
f. Attended the 8th Int'l Conf. on Control, Automation, Robotics, and Vision, Kunming, China, Dec 6-9, 2004.
g. Attended and presented papers at SPIE Watermarking and Multimedia Security Conference, San Jose, 2005.
h. Attended and presented papers at the IEEE Int'l Conf on Acoustics, Speech, and Signal Processing, March 18-23, 2005, Philadelphia, PA.

8. New discoveries, inventions, or patent disclosures. None.

9. Honors/Awards: None

10. Markings: None